Estimating Stellar Oscillation-Related Parameters and Their Uncertainties with the Moment Method

Joris De Ridder

Institute of Astronomy - Katholieke Universiteit Leuven, B-3001 Leuven, Belgium

Geert Molenberghs

Center for Statistics - Limburgs Universitair Centrum, B-3590 Diepenbeek, Belgium

Conny Aerts

Institute of Astronomy - Katholieke Universiteit Leuven, B-3001 Leuven, Belgium

Summary. The moment method is a well known mode identification technique in asteroseismology (where 'mode' is to be understood in an astronomical rather than in a statistical sense), which uses a time series of the first 3 moments of a spectral line to estimate the discrete oscillation mode parameters ℓ and m. The method, contrary to many other mode identification techniques, also provides estimates of other important continuous parameters such as the inclination angle α, and the rotational velocity v_e. We developed a statistical formalism for the moment method based on so-called generalized estimating equations (GEE). This formalism allows the estimation of the uncertainty of the continuous parameters taking into account that the different moments of a line profile are correlated and that the uncertainty of the observed moments also depends on the model parameters. Furthermore, we set up a procedure to take into account the mode uncertainty, i.e., the fact that often several modes (ℓ, m) can adequately describe the data. We also introduce a new lack of fit function which works at least as well as a previous discriminant function, and which in addition allows us to identify the sign of the azimuthal order m. We applied our method to the star HD181558 using several numerical methods, from which we learned that numerically solving the estimating equations is an intensive task. We report on the numerical results, from which we gain insight in the statistical uncertainties of the physical parameters involved in the moment method.

Keywords: Generalized estimating equations, time series, sandwich estimator, astrostatistics, discriminant function

1. Introduction

Stars consist of a number of gas layers with different temperatures, pressures, and chemical compositions. During their sojourn on the main-sequence, i.e., when they transform hydrogen into helium, some stars are subject to oscillations which in turn provide astronomers with a wealth of information about the stellar interior. This is the subject of asteroseismology. Such oscillations typically exhibit multiple frequencies and manifest themselves at the surface of the star through variations in brightness, temperature, and surface velocity; some of these are observable. A star can oscillate in one or more of its "natural" frequencies determined by the internal structure of the star. With suitable inversion techniques it is possible to use the observed frequencies to derive information about this internal structure.

To do so, however, the characteristics of the oscillations need to be considered first. That is, a *mode identification* (note that 'mode' is an astronomical term) has to be carried out, in

which one estimates the parameters characterising the oscillations from observational data. There are few mode identification techniques, and the properties of their estimators are rarely studied. Statistical uncertainties of the estimates, for example, are never reported. Nevertheless, from an astrophysical point of view, such uncertainties are important because wrong mode identifications can bias inversion techniques. It is therefore necessary to know a priori the extent of possible errors in the estimates.

In this paper, we study the statistical properties of one particular mode identification technique, the so-called *moment method*. For examples of applications of this method we refer to, e.g., Aerts et al. (1998), Uytterhoeven et al. (2001), Aerts & Kaye (2001), and Chadid et al. (2001).

2.   Astrophysical Background

As in any inferential method, the moment method uses a theoretical model to describe the observations. To understand the statistically relevant properties of this theoretical model and its parameters, we first briefly discuss some of the physics of stellar oscillations and how they are observed.

Figure 1 gives a diagrammatic illustration of an orthographic planar projection of the surface of an oscillating star, i.e., parallel with the line of sight. For our application, the
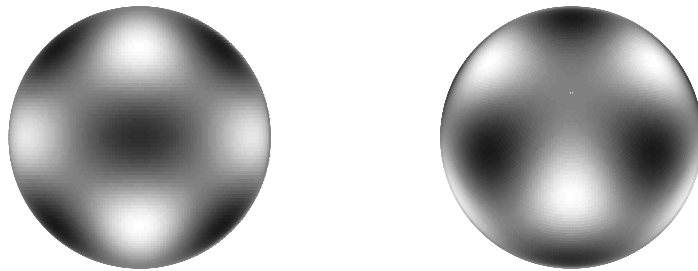

Fig. 1. A diagrammatic illustration of an orthographic planar projection of the surface of an oscillating stars for the mode $(\ell, m) = (5, 3)$. In the left picture we look at the equator, and in the right picture we look almost at the pole of the star.

most important aspect of stellar oscillation is the surface velocity. The lighter parts of the stellar surface have an inward velocity while the darker parts have an outward velocity. The figure is only a snapshot: the star varies periodically and half an oscillation cycle later the situation is reversed with the lighter parts moving outward and the darker parts moving inward. For slowly-rotating oscillating stars, each of the oscillation modes can be described with a spherical harmonic Y_ℓ^m (which is actually the basis of our illustration in Figure 1), where ℓ is the total number of nodal lines and m is the number of nodal lines perpendicular to the equator. In reality, the motion of a surface element is more complex because it moves horizontally as well as vertically.

In terms of model parameters we have to estimate 3 unknown parameters per oscillation mode: 2 discrete parameters and 1 continuous parameter. The 2 discrete parameters are the mode numbers ℓ and m of the spherical harmonic, which describe the configuration of the inward and outward going regions. To describe the 3-dimensional motion of the stellar matter, only one parameter is needed: the amplitude v_p of the vertical motion, since there is a theoretical linear relation between the amplitude of the vertical motion and the amplitude

Table 1. A summary of all unknown relevant model parameters, with their meaning and their physical range. The notation $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)'$ for the continuous parameters will be introduced and used in Section 3.

Parameter	Meaning	Physical Range
ℓ	Degree of the spherical harmonic Y_ℓ^m	$\{0, 1, 2, \cdots\}$
m	Azimuthal order of the spherical harmonic Y_ℓ^m	$\{-\ell, \cdots, 0, \cdots, +\ell\}$
$v_p = \beta_1$	Velocity amplitude of the oscillation	≥ 0
$\sigma = \beta_2$	Width of line profile in absence of pulsation and rotation (nuisance)	≥ 0
$v_e = \beta_3$	Equatorial rotational velocity	> 0
$\alpha = \beta_4$	Inclination angle of the star	$[0°, 360°)$

of the horizontal motion. To compute the constant of proportion K, however, the mass and the radius of the star are required and these quantities are often not very accurately known. Nevertheless, in what follows we will assume, as a first approximation, that K is known, to considerably simplify the treatment.

A further continuous parameter related to the oscillation is the oscillation period P. However, for good datasets, this oscillation period can often be quite accurately determined from the data with other methods so that it is usually regarded as known.

In the model, 2 additional unknown parameters not connected to the oscillations are present. A first one is the rotational velocity at the equator of the star, usually denoted as v_e. The second one is the inclination angle α under which we observe the star. This is illustrated in Figure 1. Both pictures show the same Y_ℓ^m, but on the left hand side we are looking on the equator, while on the right hand side we are looking almost on the pole. Clearly, α has a large impact on how the surface velocity field is observed.

A last unknown model parameter is specifically related to the kind of observational data we use. In the case of the moment method it concerns high-resolution spectroscopic data. The gathered star light is decomposed into its colours so that a detailed spectrum can be constructed, i.e., received light flux as a function of the wavelength of the light. At certain wavelengths, such a spectrum contains absorption lines where the light has been partially blocked by certain chemical elements at the surface of the star. An example of the Si^+ absorption line at $\lambda = 412.805$ nm for the non-radially oscillating star HD181558 is shown in the left hand panel of Figure 2. Here, an observational time series of $N = 30$ high-quality spectra gathered by De Cat and Aerts (2002) is shown. The oscillations in the star cause the absorption line to change its position and shape in time. Precisely these line profile variations are used to estimate the parameters mentioned above. To model them, another unknown parameter is needed, denoted by σ, which is related to the width the line profile would have in the absence of pulsation. From an astrophysical point of view this is an unimportant nuisance parameter. For convenience, Table 1 summarizes all unknown model parameters mentioned above, their meaning and their physical range.

Modeling the line profiles themselves turns out to be *very* computationally expensive. That is why Balona (1986) devised the moment method, which replaces each line profile by the first, second, and third moment denoted by y_1, y_2, and y_3 respectively. These quantities are *measures* for the average position, the square of the width and the skewness of the line
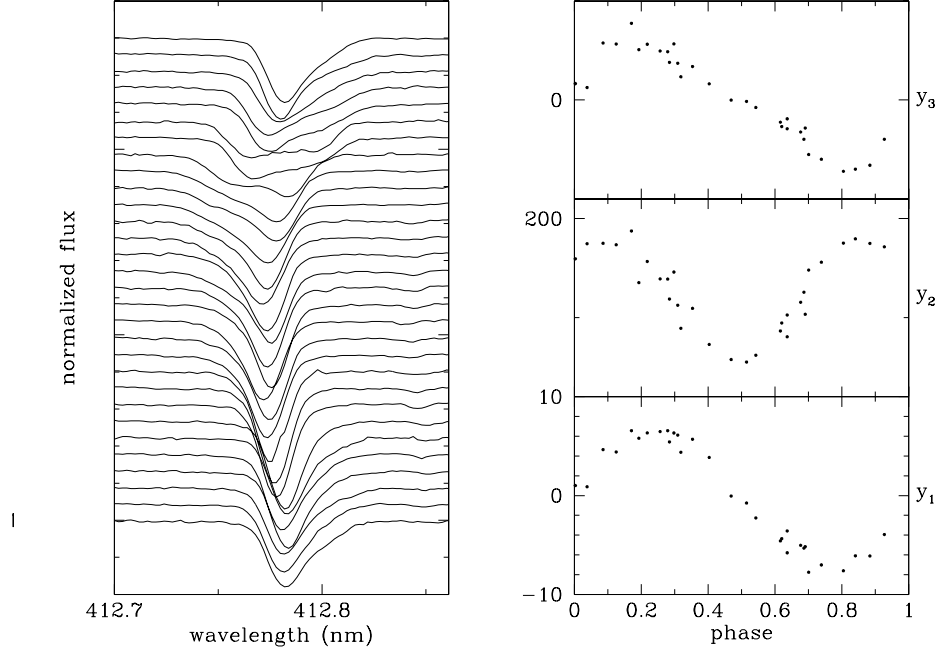
Fig. 2. In the left panel, a time series of Si^+ (412.805 nm) absorption lines of the non-radially oscillating star HD181558 is shown. The line profiles are 'sorted' to cover an entire oscillation cycle of the dominant mode, which has a period of about 29h42m. Each of the line profiles is vertically shifted to obtain a clear visual effect. In the right panels, the first moment y_1 (in km/s), the second moment y_2 (in km^2/s^2) and the third moment y_3 (in km^3/s^3) of all line profiles are shown as a function of the oscillation phase ϕ.

profile. Precisely,

$$y_n(\phi) = \frac{\int_{-\infty}^{+\infty}[1 - p(\phi, \lambda)]\lambda^n \, d\lambda}{\int_{-\infty}^{+\infty}[1 - p(\phi, \lambda)] \, d\lambda}, \tag{1}$$

where $p(\phi, \lambda)$ is the line profile function at phase ϕ and for wavelength λ. For each ϕ, there is a separate line profile in the left hand panel of Figure 2, leading to a point for each of the right hand side sub-panels, corresponding to $n = 1, 2, 3$. In practice, no higher moments are considered since these are often noisy and unduly complicate the calculations. One commonly expresses the moments in $(km/s)^n$, by transforming the wavelength λ in (1) to a velocity using the Doppler transformation formula:

$$v = c\frac{\lambda - \lambda_0}{\lambda_0},$$

where c is the speed of light. The moments $y_n(\phi)$ can be expressed in terms of time t as well, where ϕ is defined as $t \bmod P/P$, with 'mod' standing for the decimal part and P the oscillation period. A time series of theoretical moments can be computed much faster than one of theoretical line profiles. The nuisance parameter σ, however, remains. In the

right hand panels of Figure 2 we show a time series of the three moments for the star HD181558. The computation of such moments from spectral line profiles take the form of intensity-weighted sums, sums of squares, and sums of cubes.

Fifteen years after the introduction of the moment method, this mode identification technique is still very relevant (see the recent references given in Section 1). Indeed, the effort required for direct computation of time series of line profiles is currently still too computationally demanding to be useful for mode identification, even under simplifying assumptions concerning the shape of the absorption line.

A theoretical moment at one point of time is computed by integrating over the contributions of all points on the visible stellar surface. Closed-form expressions for the moments exist (Aerts et al., 1992), but they are quite lengthy and of little practical use to computing derivatives, such as in Section 5. We opt for different, computationally more advantageous, expressions, which involve an integration with bounds depending on the inclination angle α.

3. Current Statistical Status

The moment method is a multi-response problem where a time series of 3 responses is used to extract 6 (i.e., 2 discrete and 4 continuous) parameters. In what follows we will use the notation

$$\mathbf{Y}_i \equiv (y_1(t_i),\ y_2(t_i),\ y_3(t_i))'$$

and

$$\boldsymbol{\mu}_i \equiv (\mu_1(t_i, \ell, m, \boldsymbol{\beta}),\ \mu_2(t_i, \ell, m, \boldsymbol{\beta}),\ \mu_3(t_i, \ell, m, \boldsymbol{\beta}))'$$

for the first three observed and theoretical moments respectively, at time point t_i ($i = 1, \cdots, n$), where $\boldsymbol{\beta} \equiv (v_p, \sigma, v_e, \alpha)'$.

It is important to understand how the moment method is currently used. Theoretically, it can be shown (Aerts et al., 1992) that for a monoperiodic star, the time dependence of the moments takes the following form:

$$\left.\begin{array}{rcl}
\mu_1 &=& a_1\ \sin(2\pi\nu t + \kappa_1), \\
\mu_2 &=& b_0 + b_1\ \sin(2\pi\nu t + \delta_1) + b_2\ \sin(4\pi\nu t + \delta_2), \\
\mu_3 &=& c_1\ \sin(2\pi\nu t + \gamma_1) + c_2\ \sin(4\pi\nu t + \gamma_2) + c_3\ \sin(6\pi\nu t + \gamma_3),
\end{array}\right\} \quad (2)$$

where ν is the oscillation frequency. The phases α_1, δ_i, γ_j are constants, while the positive amplitudes a_1, b_i, c_j depend on the parameters $(\ell, m, \boldsymbol{\beta})$. A *discriminant* $\Gamma_l^m(\boldsymbol{\beta})$ is constructed to estimate these parameters by comparing the observed amplitudes with their theoretical counterparts:

$$\Gamma_\ell^m = \left\{ \left(f_{\tilde{a}_1}\ |\tilde{a}_1 - a_1| \right)^2 + \sum_{i=0}^{2} \left(f_{\tilde{b}_i}\ \sqrt{|\tilde{b}_i - b_i|} \right)^2 + \sum_{i=1}^{3} \left(f_{\tilde{c}_i}\ \sqrt[3]{|\tilde{c}_i - c_i|} \right)^2 \right\}^{1/2}, \quad (3)$$

where the tilde denotes observed quantities, and where the weights f are introduced to incorporate the estimated standard errors $\Delta\tilde{a}_1$, $\Delta\tilde{b}_i$ and $\Delta\tilde{c}_j$ of the corresponding observed amplitudes:

$$f_{\tilde{a}_1} \equiv W^{-1}\ \frac{\tilde{a}_1}{\Delta\tilde{a}_1}, \qquad f_{\tilde{b}_i} \equiv W^{-1}\ \frac{\tilde{b}_i}{\Delta\tilde{b}_i}, \qquad f_{\tilde{c}_i} \equiv W^{-1}\ \frac{\tilde{c}_i}{\Delta\tilde{c}_i},$$

$$W \equiv \frac{\tilde{a}_1}{\Delta \tilde{a}_1} + \frac{\tilde{b}_0}{\Delta \tilde{b}_0} + \frac{\tilde{b}_1}{\Delta \tilde{b}_1} + \frac{\tilde{b}_2}{\Delta \tilde{b}_2} + \frac{\tilde{c}_1}{\Delta \tilde{c}_1} + \frac{\tilde{c}_2}{\Delta \tilde{c}_2} + \frac{\tilde{c}_3}{\Delta \tilde{c}_3}$$

(Aerts 1996). The form of Γ_ℓ^m in (3) prevents the third moment y_3 with its large values from dominating the first moment y_1, but has the disadvantage that it cannot discriminate the sign of the mode number m. The parameters are estimated by searching for the minimum of Γ_ℓ^m in a rectangular grid in the parameter space. For the continuous parameters, it is hoped for that the grid is fine enough in order not to miss the global minimum. Finally, a table is produced with the top 5 or 6 best fitting $(\ell, m, \boldsymbol{\beta})$ parameter sets.

The best strategy to obtain the final estimate for (ℓ, m) and $\boldsymbol{\beta}$ together with their uncertainties, is currently open to debate. Despite the usefulness of the table with the best parameter sets, there is no estimate of the uncertainties of the parameters obtained with the moment method, because of severe theoretical and computational complexity. This paper takes an important first step towards estimating the uncertainties of the continuous parameters $\boldsymbol{\beta}$. Estimating the uncertainties of the discrete parameters ℓ and m is an even more challenging problem, and will be left for future research.

Even for a given (ℓ, m) value, it is currently unknown how precise the continuous parameters are estimated. For example, is the uncertainty in the inclination angle as small as $5°$, or is perhaps $30°$ a more typical value? Moreover, very often several (ℓ, m) pairs give almost equally good fits. The question is raised as to how should we take this into account for our best estimate of $\boldsymbol{\beta}$ and its uncertainty? In what follows, we will try to answer these questions.

4.   New Statistical Approach

We consider a new estimating method which produces both point and interval estimates.

We first note that the three responses y_1, y_2 and y_3 are dependent, and that their covariance matrix $\mathbf{V}$ is unknown. Formulating a statistical model for the noise on the moments is non-trivial as it would involve a model for both the instrumental and the atmospheric noise. In addition, we note that the relation between the coefficients a_1, b_i, c_i in (2), and the parameters $\boldsymbol{\beta}$ is non-linear thereby preventing the easy computation of a Jacobian matrix. Estimating $\boldsymbol{\beta}$ and its covariance matrix $\text{Cov}[\boldsymbol{\beta}]$ using a simple variable transformation technique is therefore not possible.

A first alternative is the least squares method. Although multi-response least-squares estimation has been used before to deal with correlated responses where the covariance matrix has to be estimated, Seber and Wild (1989) show that this technique should *not* be used if the covariance matrix $\mathbf{V}$ depends on the parameters $\boldsymbol{\beta}$, as is the case here. For example, the uncertainty in the first moment of a line profile (y_1) can be estimated with the second moment (y_2), and the latter depends on ℓ, m, and $\boldsymbol{\beta}$. Or, with an astrophysical example, the faster the star rotates (larger v_e) the broader and flatter the line profile, and the less precision with which we know the position or the first moment of the line profile. To avoid confusion, we stress that our argument in the example above is not that the uncertainty of the first moment y_1 depends on the uncertainty of the second moment y_2, but that the uncertainty of the first moment y_1 always depends on the second moment itself of which we know that it depends in turn on the parameters $\boldsymbol{\beta}$.

We must therefore conclude that in the case of the moment method, minimizing the weighted sum of squares is not appropriate, regardless of how $\mathbf{V}$ is estimated, because it

will not yield a consistent estimate of the parameters $\boldsymbol{\beta}$ and it will reduce the efficiency of the estimator (Seber and Wild, 1989).

The *generalized estimating equations* (GEE) methodology, as developed by Liang and Zeger (1986), is better suited for the purpose of the moment method. We recall that this method does not assume a particular *joint* probability density for the responses y_1, y_2 and y_3, nor that they are i.i.d. The theory does not assume that the theoretical model is linear in its parameters, while the covariance matrix $\mathbf{V}$ of the responses does not need to be completely specified. The method *does* assume, however, that the different observations $\mathbf{Y}_i$ $(i = 1, \cdots, n)$ are independent, that a working approximation of the covariance matrix of the responses is available, and that the expectation values $E[\mathbf{Y}_i] \equiv \boldsymbol{\mu}_i(\ell, m, \boldsymbol{\beta})$ $(i = 1, \cdots, n)$ are correctly specified.

We therefore use GEE to estimate the uncertainties of the continuous parameters $\boldsymbol{\beta}$. We recall that in the GEE method, the parameters are estimated by locating the root of the quasi-score function $\mathbf{U}(\boldsymbol{\beta})$:

$$\mathbf{U}(\boldsymbol{\beta}) \equiv \sum_{i=1}^{N} \mathbf{D}_i^t \cdot \mathbf{W}_i^{-1} \cdot (\mathbf{Y}_i - \boldsymbol{\mu}_i), \tag{4}$$

where N is the size of the time series. The 3×4 matrix $\mathbf{D} = \partial\boldsymbol{\mu}/\partial\boldsymbol{\beta}^t$, and the 3×3 symmetric matrix $\mathbf{W}_i$ is a working approximation of the true covariance matrix $\mathbf{V}_i$ of the quantities $\mathbf{Y}_i$:

$$\mathbf{V}_i \equiv E[(\mathbf{Y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta}))(\mathbf{Y}_i - \boldsymbol{\mu}_i(\boldsymbol{\beta}))^t], \tag{5}$$

where $\boldsymbol{\beta}$ are the true (but unknown) parameters. It can be shown (e.g., Liang and Zeger, 1986; Zeger and Liang 1986; Diggle et al. 2002) that the root $\hat{\boldsymbol{\beta}}$ is a consistent and asymptotically normal estimate of the true $\boldsymbol{\beta}$, with sandwich covariance matrix

$$\mathrm{Cov}[\hat{\boldsymbol{\beta}}] = \mathbf{I}_0^{-1} \ \mathbf{I}_1 \ \mathbf{I}_0^{-1}, \tag{6}$$

where

$$\mathbf{I}_0 \equiv -E\left[\frac{\partial\mathbf{U}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}}\right] = \sum_{i=1}^{N} \mathbf{D}_i^t \ \mathbf{W}_i^{-1} \ \mathbf{D}_i$$

and

$$\mathbf{I}_1 \equiv \mathrm{Cov}[\mathbf{U}(\boldsymbol{\beta})] = \sum_{i=1}^{N} \mathbf{D}_i^t \ \mathbf{W}_i^{-1} \ \mathbf{V}_i \ \mathbf{W}_i^{-1} \ \mathbf{D}_i.$$

The unknown covariance matrices $\mathbf{V}_i$ in the expression for $\mathbf{I}_1$ are estimated by

$$\hat{\mathbf{V}}_i = (\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}})) \cdot (\mathbf{Y}_i - \boldsymbol{\mu}_i(\hat{\boldsymbol{\beta}}))^t.$$

The so-called sandwich estimator in (6) is robust against misspecification of the covariance matrix of the responses.

For the working approximation $\mathbf{W}_i$ for the covariance matrix $\mathbf{V}_i$, we suggest the following idea, where we estimate the uncertainty of the first three moments of the line profile with the higher moments, as is sometimes done with the moments of a probability distribution function. Consider the mirror image $\zeta(\phi, v) = 1 - p(\phi, v)$ of the spectral line $p(\phi, v)$

as a distribution function for the velocity v, and compute, for a given time t_i:

$$
\begin{aligned}
W_{i,rs} \equiv W_{rs} &= E[(y_r - \mu_r)(y_s - \mu_s)] \\
&= E\left[\left(\frac{\sum \zeta_j \, v_j^r}{\sum \zeta_j} - \mu_r\right)\left(\frac{\sum \zeta_j \, v_j^s}{\sum \zeta_j} - \mu_s\right)\right] \\
&= \frac{1}{(\sum \zeta_j)^2} \, E\left[\sum_j \zeta_j(v_j^r - \mu_r) \cdot \sum_j \zeta_j(v_j^s - \mu_s)\right] \\
&= \frac{1}{(\sum \zeta_j)^2} \sum_j \zeta_j^2 \, (E[v_j^r \, v_j^s] - \mu_r \, \mu_s) \\
&= \Gamma \cdot (\mu_{r+s} - \mu_r \, \mu_s),
\end{aligned}
$$

where the sum over the index j runs over all the velocity points (pixels) of the spectral line, and where we define

$$
\Gamma \equiv \sum_j \zeta_j^2 \left/ \left(\sum_j \zeta_j\right)^2 \right. .
$$

Here, we assume that the different observed points of the line profile are uncorrelated. The extra factor Γ appears because, contrary to probability distribution functions, line profiles are not normalized in area. Note that we use the higher *theoretical* moments and not the observational ones because, as mentioned before, the latter are often too noisy. It is difficult to assess the influence of the working approximation $\mathbf{W}$ on the final uncertainties on $\boldsymbol{\beta}$, but we refer to Diggle et al. (2002) where it is shown that the sandwich estimator (6) for the covariance matrix of $\boldsymbol{\beta}$ is quite robust against misspecification of $\mathbf{V}$.

Having derived an estimator for $\boldsymbol{\beta}$ and its uncertainty, given an (ℓ, m) pair, we should take into account that we do not actually know the correct (ℓ, m) values. If ℓ and m were continuous parameters, we would have a total of 6 continuous parameters for which we would have liked to compute a 6-dimensional confidence region. As ℓ and m are discrete, however, it is notoriously hard to find an equivalent "confidence region". We remind that it is current practice simply to take the $\boldsymbol{\beta}$ values of the best-fitting (ℓ, m) pair with no error estimate at all. As a first alternative, we propose to "weight" each mode (ℓ, m) with a lack-of-fit function. The best guess for both $\boldsymbol{\beta}$ and its uncertainty is then computed with a weighted mean over all relevant modes (ℓ, m). The entire estimation procedure can be summarized as follows:

(a) Specify a set of pairs of the degree ℓ and the azimuthal number m: $\{(\ell_j, m_j)\}$.

(b) For each of the pairs (ℓ_j, m_j), solve the quasi-score equations and estimate the continuous parameters $\hat{\boldsymbol{\beta}}_j$ and their covariance matrix $\mathrm{Cov}[\hat{\boldsymbol{\beta}}_j]$.

(c) Compute for each of the modes (ℓ_j, m_j), the lack-of-fit parameter G_j^2 which indicates how well the theoretical moments $\boldsymbol{\mu}(\hat{\boldsymbol{\beta}})$ fit the observed moments $\mathbf{y}$:

$$
G_j^2 = \sum_{k=1}^{3} \sum_{i=1}^{N} \frac{(y_k(t_i) - \mu_k(\hat{\boldsymbol{\beta}}_j, t_i))^2}{\mu_{2k}(\hat{\boldsymbol{\beta}}_j, t_i) - \mu_k^2(\hat{\boldsymbol{\beta}}_j, t_i)}. \tag{7}
$$

(d) The best estimate for the degree and the azimuthal number $(\tilde{\ell}, \tilde{m})$ is the (ℓ_j, m_j) that has the lowest lack-of-fit G_j^2. The corresponding best estimate for the continuous

parameters $\tilde{\boldsymbol{\beta}}$ can be computed with

$$\tilde{\boldsymbol{\beta}} = \frac{\displaystyle\sum_{\{(\ell_j, m_j)\}} \hat{\boldsymbol{\beta}}_j \, G_j^{-2}}{\displaystyle\sum_{\{(\ell_j, m_j)\}} G_j^{-2}} \tag{8}$$

and the corresponding covariance matrix is the sum of the intra-mode variance and the inter-mode variance:

$$\mathrm{Cov}[\tilde{\boldsymbol{\beta}}] = \frac{\displaystyle\sum_{\{(\ell_j, m_j)\}} \mathrm{Cov}[\hat{\boldsymbol{\beta}}_j] \, G_j^{-2}}{\displaystyle\sum_{\{(\ell_j, m_j)\}} G_j^{-2}} + \frac{\displaystyle\sum_{\{(\ell_j, m_j)\}} (\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_j) \cdot (\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}_j)^t \, G_j^{-2}}{\displaystyle\sum_{\{(\ell_j, m_j)\}} G_j^{-2}}. \tag{9}$$

For both practical and astrophysical reasons, only modes with a degree ℓ up to a certain limit (e.g. $\ell \leq 4$) are considered.

In the following section this estimation procedure is applied to a dataset of the star HD181558.

5.   Application to HD181558

HD181558 belongs to the class of the Slowly Pulsating B stars (SPBs). Although the star is multi-periodic (De Cat and Aerts 2002) it has a very dominant (in amplitude) first mode, which justifies a monoperiodic approximation. The amplitude of this mode is the largest ever observed for an SPB. The dataset used for this GEE application has already been shown in Figure 2. In what follows we always assume the theoretically predicted value $K = 21$.

Our first goal was to estimate $\boldsymbol{\beta}$ for each mode (ℓ, m) with $\ell \leq 4$, by solving the non-linear quasi-score equations. It turned out, however, that this was not just a technical detail of the procedure, but was in fact a major issue.

First, it turned out that the quasi-score function $\mathbf{U}(\boldsymbol{\beta})$ is computationally slow to evaluate, with one evaluation requiring 18 time series evaluations: 6 for the moments μ_1–μ_6 for the working approximation $\mathbf{W}$, and 12 for the moments μ_1–μ_3 for different parameters $\boldsymbol{\beta}$ to numerically compute (with forward differences) the derivatives in $\mathbf{D}$. For this reason, prior to using (4), we first determined a good initial guess for $\hat{\boldsymbol{\beta}}$ for the local search routine, using a rough scan of the 4D parameters space for each mode (ℓ, m) with a computationally less expensive lack-of-fit function $g(\boldsymbol{\beta})$:

$$g(\boldsymbol{\beta}) \equiv \sum_{d=1}^{3} \frac{1}{d} \sqrt[d]{\frac{1}{N} \sum_{i=1}^{N} |y_d(t_i) - \mu_d(t_i, \boldsymbol{\beta})|}. \tag{10}$$

The construction with the dth root and the division by d simply prevents the higher order moments from numerically dominating the lower order moments. The sampling of the parameter space was done probabilistically and non-uniformly. For each parameter β_i, a physical range was determined and this range was subdivided into intervals. After each set of 10000 sampled points, each interval of each parameter β_i was assigned a sampling probability according to the lowest $g(\boldsymbol{\beta})$ value recorded up to then, with the β_i component

Table 2. For each (ℓ, m) pair, the 4D parameter space was scanned with the lack-of-fit function g defined by equation (10) and with the dataset of the star HD181558 shown in Figure 2. The minimum $g_{\min}$ of the $g(\boldsymbol{\beta})$ values for each of the (ℓ, m) pairs is given.

		ℓ		
m	1	2	3	4
+4				11.9
+3			7.52	11.0
+2		6.57	6.71	11.3
+1	4.72	4.74	5.85	10.8
0	6.37	6.37	6.79	11.6
−1	4.79	6.57	7.05	10.7
−2		4.68	6.86	10.5
−3			6.92	11.3
−4				11.3

in the corresponding interval. The sum of probabilities over all intervals of a parameter β_i was set to one. For each (ℓ, m) pair a total of 200,000 points was sampled. This procedure was set up to sample the more promising regions of the parameter space.

In Table 2 we give for the star HD181558 the lowest $g(\boldsymbol{\beta})$ value recorded for each mode (ℓ, m). The $(\ell, m) = (0, 0)$ pair can be excluded on astrophysical grounds because such modes do not occur in SPBs. As can be seen, no mode (ℓ, m) stands out, but there are several candidate modes that describe the data well. Our final estimate of $\boldsymbol{\beta}$ should take into account this mode uncertainty. We also note that Table 2 is not symmetric with respect to the sign of m. The moments indeed behave differently when the pulsational wave goes in the same direction as the rotation then when the wave goes in the opposite direction. This sensitivity to the sign of m was lost in the old approach (outlined in Section 3) where one only uses the absolute value of the amplitudes.

The 24 scans of a 4D parameter space with 200,000 points each, was a rather time consuming but necessary task to find suitable initial guesses for $\hat{\boldsymbol{\beta}}$ for the local search algorithm. We implemented two derivative-free methods: the conjugate-direction (Powell's) method (see, e.g., Press et al., 1992, p. 420) and the Torczon (1989) simplex method. The former of which, having the best performance, was used to locate the root $\hat{\boldsymbol{\beta}}$ of $\mathbf{U}$ for all modes (ℓ, m). We found that these methods had much stabler performance than quasi-Newton methods, such as Newton-Raphsons, Fisher scoring, or variations to this theme.

Even with the conjugate-direction method, the algorithm did not always converge. The reason, as it turns out, is that the quasi-score functions have "false" zeros, for example there are cases where the components of $\mathbf{U}$ approach zero for $\sigma \to \infty$. Quite often, the algorithm converged to a point outside the physically relevant range of the parameters, even when several different initial guesses for $\hat{\boldsymbol{\beta}}$ were tried. Although they did *not* occur for our dataset of the star HD181558, we should mention two other possible causes of numerical difficulties. First, it may be possible that the working approximation $\mathbf{W}$ is not invertible, for example if $\boldsymbol{\beta}$ approaches zero. Second, the matrix $\mathbf{I}_0$ may not be invertible, and hence no covariance

matrix can be computed. This occurs, for example, for $\alpha \rightarrow 0°$ because v_e appears only in $v_e \sin \alpha$ in the equations, so that the third row and the third column of $\mathbf{I}_0$ are zero. We stress, however, that the latter example is a problem of intrinsic non-identifiability and is not specific for the GEE approach. One simply cannot derive the rotational velocity if the star is looked pole-on.

Making detailed 1D slices of the 4D function $\|\mathbf{U}\|$ too time consuming, but we record the minimal lack-of-fit values $g_{\min}$ in each of the intervals of each parameter (disregarding the values of the other parameters). Figure 3 shows typical examples. Although the function
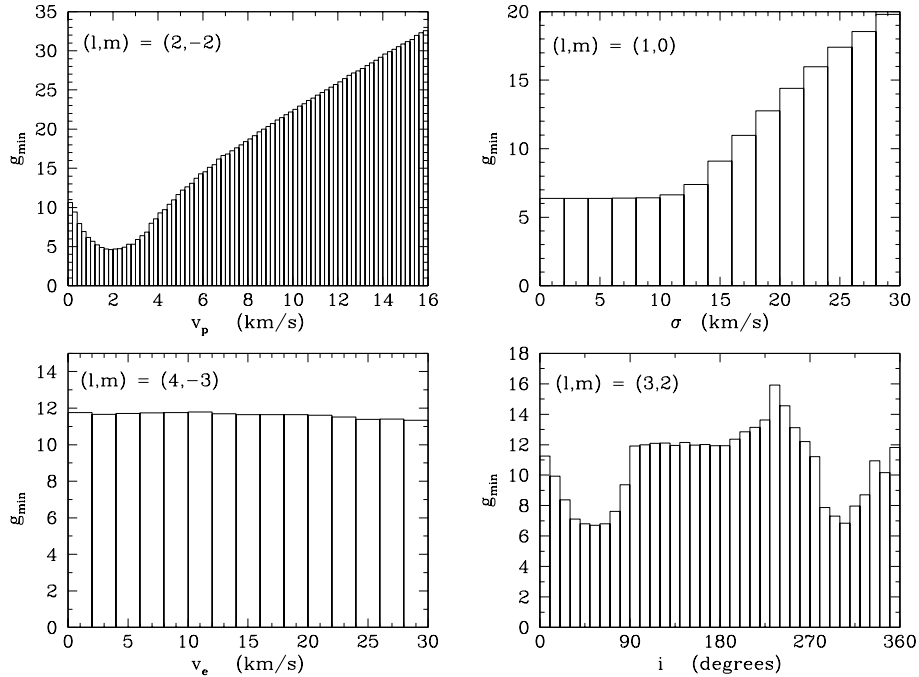

Fig. 3. Representative examples of the minimal lack-of-fit value $g_{\min}$ for each sample interval of a parameter, for the star HD181558. We remark that although the size of the intervals for the parameter v_p is fixed, the relevant range of v_p depends on the mode numbers (ℓ, m).

$g(\boldsymbol{\beta})$ need not have exactly the same behaviour as the function $\|\mathbf{U}(\boldsymbol{\beta})\|$ (the difference is similar to the well-known difference between L_1-norm and L_2-norm minimization), we assume that the functions share many features. We observe that the minimum in the upper left panel of Figure 3 for the well-fitting mode $(\ell, m) = (2, -2)$ is quite localized. This is much in contrast with the almost flat surface in the lower left panel for the badly-fitting mode $(\ell, m) = (4, -3)$. Intuitively, one can expect that the equivalent for the case of the $\|\mathbf{U}\|$ function hampers the iterations towards the minimum, and that this increases the chance of wandering from the physically relevant part of the parameter space. This is exactly what happened for this mode. More generally, we observe strong correlation between how well a mode fits the data (with $g_{\min}$ as lack-of-fit value) and the chances that the root finding algorithm does not converge. The lower right panel shows two minima in the plot of the inclination angle α, since we consider a full 360° range. While over such a range symmetry

Table 3. Roots of the quasi-score functions for those modes where there was convergence in the physically relevant part of the parameter space, for the star HD181558. The values between brackets are the standard errors obtained with the sandwich estimator (6). G^2 is the lack-of-fit value of the mode as defined by Eq. (7). v_p, σ and v_e are expressed in km/s, and the inclination angle α in degrees.

(ℓ, m)	$\|\mathbf{U}\|_{\min}$	G^2	$\widehat{v_p}$	$\widehat{\sigma}$	$\widehat{v_e}$	$\widehat{\alpha}$
(1,0)	0.15	2.7	2 (1)	9.0 (0.9)	13 (19)	320 (48)
(1,1)	$5.1\ 10^{-24}$	0.63	2.01 (0.08)	6.3 (0.2)	15.2 (0.7)	117 (2)
(1,-1)	$1.1\ 10^{-5}$	1.1	4.0 (0.2)	4.2 (0.6)	25 (2)	336 (1)
(2,0)	0.10	3.1	0.9 (0.4)	7 (2)	30 (24)	331 (13)
(2,1)	1.7	0.61	1.7 (0.1)	3.8 (0.8)	16 (1)	71 (1)
(2,2)	0.014	2.4	1 (1)	10 (2)	0.4 (29)	270 (360)
(2,-2)	0.0032	0.72	1.62 (0.06)	4.3 (0.4)	17.6 (0.7)	129 (1)
(3,1)	$1.1\ 10^{-25}$	2.3	1.00 (0.04)	6.8 (0.7)	18 (2)	145 (7)
(3,2)	0.40	3.1	1.1 (0.2)	6 (3)	17 (10)	49 (23)
(3,-1)	3.5	7.0	1.3 (0.3)	3 (12)	7 (12)	189 (5)
(4,0)	2.7	11	0.4 (0.3)	8 (12)	49 (188)	19 (96)
(4,-4)	0.047	8.7	0.6 (3)	7 (24)	20 (87)	295 (30)

relations exist, these depend on ℓ and m and taking them into account to limit the range of α would entail a lot of bookkeeping that can elegantly be avoided by simply considering the entire range. The feature that we quite often do not seem to find the root of $\mathbf{U}$ does not necessarily contradict the theory outlined in Section 4. We solve for the root of the observed $\mathbf{U}$ function because we know that $E[\mathbf{U}] = \mathbf{0}$. However, the latter is only true if the model is correctly specified, i.e. if $E[\mathbf{Y}] = \boldsymbol{\mu}(\ell, m, \boldsymbol{\beta})$. Therefore, theoretically, the existence of a root in the 4D parameter space of the continuous parameters $\boldsymbol{\beta}$ cannot be guaranteed for a "wrong" (ℓ, m) pair, and this is exactly what we observe for badly fitting modes. For this reason we interpreted a non-convergence (after repeatedly trying) as an indication that the candidate mode should be disregarded.

In Table 3, we list the roots of the quasi-score function for those modes for which there was convergence in the physically relevant part of the parameter space. The closeness of $\|\mathbf{U}\|$ to zero varies from mode to mode. In cases where this value fails to be small, we checked this is not due to premature convergence, since restarting the algorithm at the point where it stopped, did not further decrease $\|\mathbf{U}\|$. One possible explanation might be that, for some of the solutions, the algorithm has converged to a local minimum, such as for the modes $(\ell, m) = (3, -1)$ and $(4, 0)$.

In Figure 4 we show fits for the three best fitting modes, with the function G^2 (see Eq. 7) as a lack-of-fit. As mentioned before, there is not just one, but several modes that can fit the observed data quite well. The lack-of-fit values in Table 2 provide us with an indication of the relative merits of wave number choice (ℓ, m). Of course, at this point we lack knowledge about the reference distribution of these values, unlike in classical fit statistics (e.g., likelihood-ratio based). However, similar instances exist in both a frequentist (e.g., Akaike Information Criterion) and a Bayesian context (e.g., Bayes factors). Nevertheless, we assert that these numbers, especially when supported by careful graphical inspection, are useful to narrow down substantially our uncertainty about the wavenumbers, in spite of an intrinsically complicated modelling endeavor. To this end, the last column in Figure 4 displays mode $(\ell, m) = (3, 1)$ with a substantially worse fit than the one in the first three columns of the same figure.
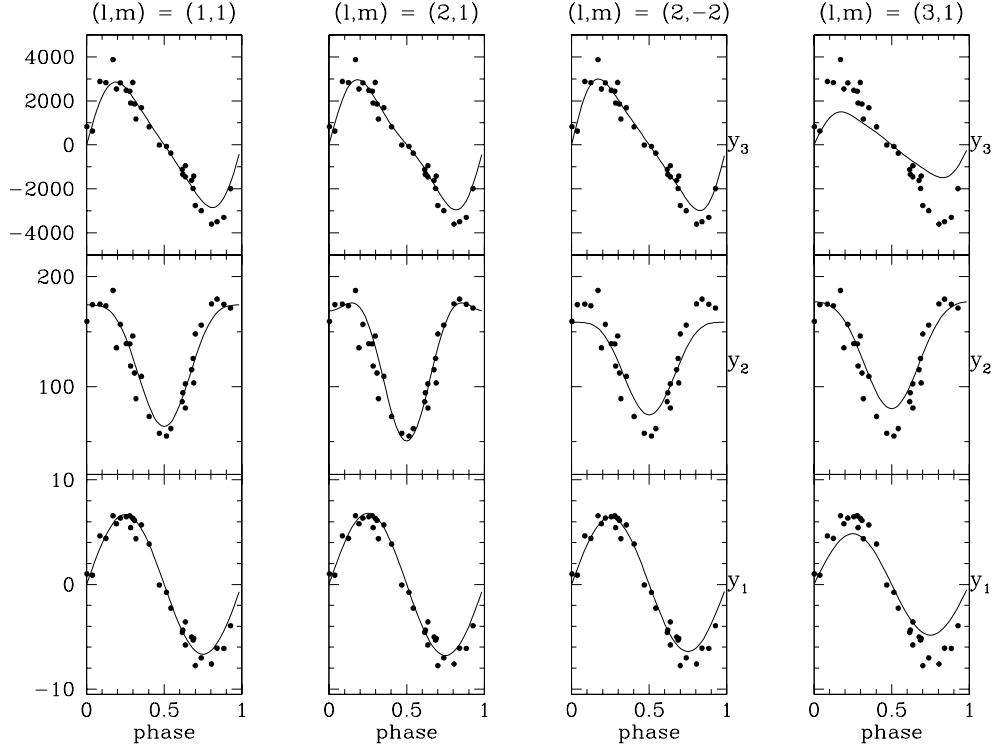
Fig. 4. Theoretical models (solid lines) of the observed moments (bullets) for the three best fitting modes $(\ell, m) = (1, 1)$, $(\ell, m) = (2, 1)$, and $(\ell, m) = (2, -2)$, plus the poorer fitting mode $(\ell, m) = (3, 1)$, with G^2 as a lack-of-fit function. The theoretical models were obtained with the model parameters obtained with the GEE method. The first, the second and the third row are for the first moment y_1 (km/s), the second moment y_2 (km^2/s^2) and the third moment y_3 (km^3/s^3) respectively. The moments are shown as a function of the phase. Note that the models of the different promising modes differ mainly for the second moment.

We used the modes in Table 3, to compute the *weighted* mean $\tilde{\boldsymbol{\beta}}$ and its standard error, with Eqs. (8) and (9). The results, including the intra-mode and inter-mode variance, are given in Table 4. In the specific case of HD181558, one could argue that the modes with $\ell = 3$ and $\ell = 4$ can be disregarded on astrophysical grounds. The reason is that these modes would require a very large oscillation amplitude at the surface of the star to cause the large observed amplitude of the first moment y_1. For this reason, we also computed $\boldsymbol{\beta}$ with the $\ell = 1$ and $\ell = 2$ modes of Table 3 only. The results are listed in Table 5. With Tables 4 and 5 we achieve the goal of this application of the revised version of the moment method: we have obtained a best guess for the continuous parameters and their standard errors, where we took into account the mode uncertainty. An important result is that the uncertainties of the parameters can be large, in fact larger than we expected. Especially the rotational velocity v_e cannot be estimated precisely. The large inter-mode uncertainty of the inclination angle α is not surprising since the inclination angle is known to be largely dependent on the mode numbers (ℓ, m). We note that the values for the weighted means

Table 4. The weighted mean over all 12 modes in Table 3, computed with Eqs. (8) and (9). The values mentioned between brackets are standard errors. The intra-mode variance and inter-mode variance are computed with respectively the first and the second term of (9). $\tilde{v}_p$, $\tilde{\sigma}$, $\tilde{v}_e$ are expressed in km/s, and the inclination angle $\tilde{\alpha}$ is expressed in degrees.

$\tilde{\beta}_i$	Weighted Mean	Intra-mode Variance	Inter-mode Variance
$\tilde{v}_p$	1.8 (1.0)	0.33	0.74
$\tilde{\sigma}$	5.5 (4.1)	14	3.1
$\tilde{v}_e$	17 (26)	612	46
$\tilde{\alpha}$	164 (132)	7300	10079

Table 5. The same information as in Table 4 is shown, except that the mean is computed over those 7 modes in Table 3 with $\ell = 1$ and $\ell = 2$.

$\tilde{\beta}_i$	Weighted Mean	Intra-mode Variance	Inter-mode Variance
$\tilde{v}_p$	2.0 (1.0)	0.19	0.71
$\tilde{\sigma}$	5.3 (2.0)	0.82	3.1
$\tilde{v}_e$	16 (12)	100	37
$\tilde{\alpha}$	170 (137)	8342	10450

do not change much by excluding the $\ell = 3$ and $\ell = 4$ modes. The reason is that the latter modes have a lower weight anyway, as can be seen from the G^2 values in Table 3.

The fact that the standard errors for the continuous parameters are large turns out not to be specific for the star HD181558. We applied our method to several artificial datasets, and we obtained similar results. Hence, we conclude that estimates of the continuous parameters generally can be quite uncertain.

6.   Summary and Conclusions

We made a first but arguably important step to develop a statistical formalism for the moment method. In this first stage we aimed to incorporate estimates of the uncertainties of the continuous parameters. Because of the many difficulties to overcome, this was never done for the moment method, nor for any other mode identification technique.

We found that, in the specific case of the moment method, the method of least-squares does not give consistent estimates of the continuous parameters and we resort to the GEE method (Liang and Zeger 1986). This method requires a working approximation of the covariance matrix of the 3 responses, based on the higher theoretical moments. Note that the higher moments, known to be imprecise, are *not* used in the actual model. An important source of uncertainty is the fact that often not just one but several candidate modes can describe the data. We set up a separate procedure to weight each mode and to compute a weighted mean over all modes of the parameter vector and its uncertainty. To compute the latter we introduced the intra-mode and the inter-mode uncertainty.

Subsequently, we applied our procedure to the SPB star HD181558, from which we learned the strong and the weak points of our method. We found out that solving the estimating equations is a computationally demanding and tedious task: convergence of the algorithm was not evident, despite the fact that we experimented with several robust local root finding methods, of which we selected the method of conjugate directions as the most efficient one. On the other hand, we also proposed a new lack-of-fit function to scan the parameter space to obtain good initial guesses for the local search method. This

lack-of-fit function proved to be very useful on its own, as it works at least as well as the old discriminant (3) and allows in addition to discriminate between positive and negative azimuthal numbers which was one of the shortcomings of the previous discriminant. Our strategy to scan the parameter space also proved that there are several modes that can explain the dataset of HD181558. Only taking into account the very best fitting one, would therefore not be useful.

This is why we retained 12 modes as candidate modes for which an estimate of the continuous parameters $\hat{\boldsymbol{\beta}}$ can be computed, and used these estimates to obtain a best guess $\tilde{\boldsymbol{\beta}}$ for the continuous parameters plus their uncertainties, taking into account the mode uncertainty. Doing so, we discovered that the parameter uncertainties can be large, a result which was moreover confirmed in the case of artificial datasets.

Prior to our study, such large uncertainties were not anticipated. On the contrary, one sometimes assumed the uncertainties to be quite small in order to be able to apply a two-stage approach in the multiperiodic case. In such an approach the inclination angle α and the rotational velocity v_e are determined with the dominant mode, and subsequently fixed while determining the mode parameters of the other modes, to have the dimension of the parameter space reduced. Our new results show that such an approach can be very dangerous: in the case of HD181558 it can hardly be justified because of the large uncertainty on α.

The method we outlined in this paper is the very first attempt to develop a statistical formalism for the moment method. Even though there is undoubtedly room for additional work before our proposed method can be deemed widely applicable, we conclude it makes an important first step in our understanding of the uncertainties and usefulness of the continuous parameters, estimated with the moment method. It furthermore underscores that some conclusions reached in the past need to be revisited.

We conclude with several possible future improvements. First, it may be worth to investigate alternative parameterizations $\boldsymbol{\beta}'$ which, while mathematically equivalent to the original one, may improve upon the convergence and robustness properties of the algorithm. We already experimented, for example, with using $v_e \sin \alpha$ instead of v_e, and with using $v_p \sin \alpha$ and $v_p \cos \alpha$ instead of v_p and α. Second, it would be useful to extend the formalism to include the uncertainty on K. Third, it might also be interesting to use several spectral lines at the same time to improve the statistics. Including multiple modes might also improve the convergence properties. Although this would imply 3 more parameters $(l_2, m_2, v_{p,2})$ per mode, multiple modes would set more stringent restrictions on the inclination angle α which has a significant impact on the estimation of the other parameters. To further validate all of the above, more simulations are necessary. In addition, such simulations can also clarify what impact the number and the signal to noise ratio of the observational spectral lines has on the performance of the algorithm.

References

Aerts, C., De Pauw, M. and Waelkens, C. (1992) Mode identification of pulsating stars from line profile variations with the moment method. An example - The Beta Cephei star Delta Ceti, *Astronomy and Astrophysics*, **266**, 294-306.

Aerts, C. (1996) Mode identification of pulsating stars from line-profile variations with the moment method: a more accurate discriminant, *Astronomy and Astrophysics*, **314**, 115-

122

Aerts, C., De Cat, P., Cuypers, J., Becker, S.R., Mathias, P., De Mey, K., Gillet, D. and Waelkens, C. (1998) Evidence for binarity and multiperiodicity in the beta Cephei star beta Crucis, *Astronomy and Astrophysics*, **329**, 137-146

Aerts, C. and Kaye, A.B. (2001) A Spectroscopic Analysis of the gamma Doradus Star HD 207223 = HR 8330, *The Astrophysical Journal*, **553**, Issue 2, 814-822

Balona, L.A. (1986) Mode identification from line profile variations, *Monthly Notices of the Royal Astronomical Society*, **220**, 647-656

Chadid, M., De Ridder, J., Aerts, C. and Mathias, P. (2001) 20 CVn: A monoperiodic radially pulsating delta Scuti star, *Astronomy and Astrophysics*, **375**, 113-121

De Cat, P. and Aerts, C. (2002) A study of bright southern slowly pulsating B stars. II. The intrinsic frequencies, *Astronomy and Astrophysics*, **393**, 965-981

Diggle, P.J., Heagerty, P., Liang, K.Y. and Zeger, S.L. (2002) *Analysis of Longitudinal Data*, Oxford University Press, Oxford

Liang, K.Y. and Zeger, S. (1986) Longitudinal data analysis using generalized linear models, *Biometrika*, **73**, 13

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) *Numerical Recipes in C. The art of scientific computing.* (2nd ed.), Cambridge University Press

Seber, G.A.F. and Wild, C.J. (1989) *Nonlinear Regression*, Wiley series in probability and mathematical statistics, John Wiley & Sons, Inc.

Torczon, V.J. (1989) Multi-directional search: a direct search algorithm for parallel machines, PhD Thesis, Rice University, Houston (Texas, U.S.A.)

Uytterhoeven, K., Aerts, C., De Cat, P., De Mey, K., Telting, J.H., Schrijvers, C., De Ridder, J., Daems, K., Meeus, G. and Waelkens, C. (2001) Line-profile variations of the double-lined spectroscopic binary kappa Scorpii, *Astronomy and Astrophysics*, **371**, 1035-1047

Zeger, S.L., Liang, K.Y. (1986) Longitudinal data anlaysis for discrete and continuous outcomes, *Biometrics*, **42**, 121-130